DOCUMENT RESUME

ED 129 330 IR 004 117

AUTHOR Harvey, Francis A.; And Others
TITLE Evaluation of Eight "Infinity Factory" Programs; Part I: Analysis of the Eight-Show Series.
INSTITUTION Education Development Center, Inc., Newton, Mass.
PUB DATE Jun 76
NOTE 61p.; For related documents see IR 004 116-119

EDRS PRICE MF-$0.83 HC-$3.50 Plus Postage.
DESCRIPTORS Attention; Comprehension; *Educational Television; Elementary Education; *Elementary School Mathematics; Grade 3; Grade 4; Grade 5; Grade 6; Humanistic Education; Knowledge Level; *Mathematics Instruction; Media Research; Mexican Americans; *Minority Group Children; Motivation; Negro Students; Program Evaluation; Self Esteem; Spanish Americans; Spanish Culture; Student Attitudes; Summative Evaluation; Teacher Attitudes
IDENTIFIERS Blacks; *Infinity Factory; Latinos

ABSTRACT

The "Infinity Factory" television series was developed to help children ages 8 through 11 to understand the usefulness of some basic mathematics skills for everyday life. Aimed primarily at black and Latino children, the series concentrates on: the decimal number system; measurement, especially the metric system; estimation; mapping and scaling; and graphing. Throughout the series there is emphasis on creative problem solving techniques and on a positive student self-image. An evaluation of eight of the programs was conducted using 1,000 students and their teachers in 39 3rd-6th grade classes in four cities as subjects. The evaluation measured student attention, appeal of the overall programs and major segments of each program, student comprehension of story line and gains in math skills, attitudes toward math, social attitudes, and teachers' opinions of the effectiveness and usefulness of the series. This report gives a detailed analysis of the evaluation of the series taken as a whole. (JY)

****
* Documents acquired by ERIC include many informal unpublished *
* materials not available from other sources. ERIC makes every effort *
* to obtain the best copy available. Nevertheless, items of marginal *
* reproducibility are often encountered and this affects the quality *
* of the microfiche and hardcopy reproductions ERIC makes available *
* via the ERIC Document Reproduction Service (EDRS). EDRS is not *
* responsible for the quality of the original document. Reproductions *
* supplied by EDRS are the best that can be made from the original. *
****

ED129330

EVALUATION OF EIGHT "INFINITY FACTORY" PROGRAMS

PART I:

Analysis of the Eight-Show Series

Francis A. Harvey, Senior Evaluator
Barbara Quiroga
Valerie Crane
Charles L. Bottoms

U S DEPARTMENT OF HEALTH
EDUCATION & WELFARE
NATIONAL INSTITUTE OF
EDUCATION

THIS DOCUMENT HAS BEEN REPRODUCED EXACTLY AS RECEIVED FROM THE PERSON OR ORGANIZATION ORIGINATING IT. POINTS OF VIEW OR OPINIONS STATED DO NOT NECESSARILY REPRESENT OFFICIAL NATIONAL INSTITUTE OF EDUCATION POSITION OR POLICY

June, 1976

IR004117

edc education development center, inc./55 chapel st./newton, mass. 02160/(617)969-7100

PERMISSION TO REPRODUCE THIS COPYRIGHTED MATERIAL HAS BEEN GRANTED BY

Education Development Center

TO ERIC AND ORGANIZATIONS OPERATING UNDER AGREEMENTS WITH THE NATIONAL INSTITUTE OF EDUCATION. FURTHER REPRODUCTION OUTSIDE THE ERIC SYSTEM REQUIRES PERMISSION OF THE COPYRIGHT OWNER.

copyright©1976 education development center, inc.

CONTENTS

Part I Analysis of the Eight-Show Series

Section I Introduction
Section 2 Method
Section 3 Analysis of the Data
Section 4 Results
Section 5 Summary and Conclusions
Section 6 Recommendations

Part II Show-by-show Analysis (Bound Separately)

Section 1 Introduction
Section 2 Method
Section 3 Analysis of the Data
Section 4 Results, Show A
Section 5 Results, Show B
Section 6 Results, Show C
Section 7 Results, Show D
Section 8 Results, Show E
Section 9 Results, Show F
Section 10 Results, Show G
Section 11 Results, Show H
Section 12 Summary and Conclusions
Section 13 Recommendations

Appendices (Bound Separately)

Appendix A Abstracts of Part II: Show-by-Show Analyses
Shows A to H
Appendix B Pretest/Posttest Instruments
Student Attitude Form (Pretest/Posttest)
Student Math Content Form (Pretest/Posttest)
Teacher Semantic Differential (Pretest/Posttest)
Teacher Opinion Form (Posttest Only)

CONTENTS (continued)

Appendices

Appendix C Weekly Instruments
- Student Weekly Response Forms, Shows A to H
- Student Attention Forms, Shows A to H
- Teacher Weekly Questionnaire

Appendix D Student Comprehension, Math Content, and Attitudes:
- Items Selected for Subscales

Appendix E Teacher Responses:
- Items Selected for Subscales

Appendix F Student Attention:
- Frequency Data

Appendix G Student Appeal: Comparisons Among Shows
- Frequency Data and Results of Statistical Analyses

Appendix H Student Appeal: Comparisons Among Groups
- Frequency Data and Results of Statistical Analyses

Appendix I Comprehension Subscale:
- Frequency Data

Appendix J Students' Knowledge of Math Content:
- Frequency Data

Appendix K Student Attitudes:
- Frequency Data

Appendix L Teacher Responses:
- Frequency Data and Results of Statistical Analyses

Appendix M Program Guides
- Introduction to Program Guide
- Program Guides, Shows A to H

CONTENTS (continued)

Appendices

Appendix N	Descriptive Statistics, Show-by-Show Analysis: Show A
Appendix O	Descriptive Statistics, Show-by-Show Analysis: Show B
Appendix P	Descriptive Statistics, Show-by-Show Analysis: Show C
Appendix Q	Descriptive Statistics, Show-by-Show Analysis: Show D
Appendix R	Descriptive Statistics, Show-by-Show Analysis: Show E
Appendix S	Descriptive Statistics, Show-by-Show Analysis: Show F
Appendix T	Descriptive Statistics, Show-by-Show Analysis: Show G
Appendix U	Descriptive Statistics, Show-by-Show Analysis: Show H

1. INTRODUCTION

"Infinity Factory" is a television series about mathematics, people, and people using math. The series was produced by Education Development Center, Newton, Mass., under a grant from the U.S. Office of Education, ESAA, with additional start-up support from the Carnegie Corporation of New York, John and Mary R. Markle Foundation, JDR 3rd Fund, National Science Foundation, and Alfred P. Sloan Foundation. Designed for both home and classroom viewing, the series presents mathematics in a common-sense way that helps children understand the usefulness of mathematics in their own lives. The programs are for children ages 8 through 11, especially Black and Latino children. A series of 52 half-hour programs has been produced and is scheduled for broadcasting over the Public Broadcasting Service beginning in the Fall of 1976.

"Infinity Factory" mathematics concentrates on five main areas:

1. Decimal number system, including single-digit arithmetic and techniques for getting rough arithmetic answers quickly, such as rounding off;
2. Measurement, with a special emphasis on the metric system;
3. Estimation;
4. Mapping and scaling, including treatments of ratio and proportion;
5. Graphing.

Intertwined through all the mathematical areas are some useful ways to solve problems: techniques that apply to problems in many areas. These methods are presented to encourage viewers to think creatively about problems they encounter themselves.

Along with the mathematics, "Infinity Factory" addresses a set of cultural and ethnic goals that reflect the special needs of minority children in the audience. These goals include:

1. presenting positive Black and Latino role models;
2. helping each viewer to reinforce good feelings about his or her own group, and to accept people and relationships in other groups;
3. representing the inner-city environment, both for urban audiences and for suburban and rural viewers, in order to present experiences common to many members of the target audience;
4. stressing the humanistic perspectives of sharing, cooperation, equality, and self-respect.

The program follows a magazine format. Each program centers around one main math theme, usually involving two or three skills or concepts. Several short segments in each program treat this mathematical theme from different perspectives, developing the mathematics in several related ways. These segments show math at work in people's everyday lives.

The major segments of each program are:

"Scoops' Place" -- a live-action, dramatic segment about a Black family who runs a neighborhood store in New York City.

"City Flats" -- also live-action, is about a Latino family operating a bakery in East Los Angeles.

"Brownstone" segments -- a resident multi-ethnic cast of young people act in short skits that usually take place in and around an urban "Brownstone" apartment house built in a television studio.

Animation segments -- used to present certain math concepts in a humorous and direct manner.

In addition, every program features a historical segment introduced by the "Brownstone" cast. Each of these segments points out an important contribution made by a notable minority person, often in a field involving mathematics. Some programs also include "Math in the Street" interviews, which present spontaneous responses from many people to a question about mathematics.

In conjunction with the trial broadcast season of the "Infinity Factory" series during the Spring of 1976, an evaluation of eight programs was conducted.* The evaluation effort spanned a ten-week period: one week of pretesting, eight weeks of in-school viewing, and a final week of posttesting. The program was viewed in four cities in the United States. Over 1,000 students and their teachers in 39 third-to-sixth-grade classes participated in the study.

The objectives of the evaluation were:

1. to determine student attention to the programs;
2. to determine the appeal of the overall programs and of the major segments of each program to both students and teachers;
3. to determine student comprehension of the dramatic story lines;
4. to determine the degree to which the eight-program "mini-series" met its objectives in the areas of learning math content, attitudes toward math, and social attitudes;

---

* The eight programs evaluated were:

Program	Broadcast Number	Topic
A	114	Measurement of Time
B	127	Rounding Off and Approximation
C	130	Measurement of Weight
D	131	Mapping and Scaling
E	103	Graphing
F	123	Estimation of Quantity
G	129	Measurement of Weight
H	132	Mapping and Scaling

5. to determine teachers' opinions of the effectiveness of the series and its usefulness in the classroom.

This evaluation report will include two parts, Part I examines the effectiveness of the eight programs taken as a whole through statistical analyses of pretest/posttest differences, subscales based on responses over eight programs, and trends over eight programs. In particular, Part I focusses on student attention, student appeal, students' comprehension of dramatic story lines, students' knowledge of math content, students' attitudes, and teachers' attitudes.

Part II of the report presents a descriptive report on each of the eight programs in the areas of student attention, student appeal, teachers' opinions of the programs, and the number and kinds of related classroom activities. Abstracts of these reports are contained in Appendix A.

2. METHOD

Subjects

Students and teachers in 39 classes participated in the evaluation study. There were 5 classes (n = 131 students) in Lawrence, Massachusetts;* 13 classes (n = 265) in Boston, Massachusetts; 10 classes (n = 319) in Los Angeles, California; and 11 classes (n = 327) in New York, New York. There were 2 third grade classes, 15 fourth grade classes, 12 fifth grade classes, and 10 sixth grade classes. The study included a total of 562 girls and 480 boys; Table 2.1 provides a further breakdown according to ethnic group, age, and sex for all students for whom complete data were available. (Since the primary focus of this evaluation was on the effectiveness of the "Infinity Factory" programs for Black and Latino students, students who were white or other minority were pooled into the category of non-target students.)

The sample of 39 teachers was comprised of 8 Black women, 2 Latino women, 23 white women, and 4 white men. The teachers had an average of 7.6 years of teaching experience. Of the 39 teachers, 14 described themselves as infrequent users of media (television and other audio-visual materials); 11 teachers described themselves as moderate users of media; and 14 teachers described themselves as frequent users of media.

---

* A small, industrial city north of Boston.

Table 2.1

Description of Sample Population

Group	Age		
	Younger Students[a]	Older Students[b]	Total
Black Students			
Boys	122	69	191
Girls	140	93	233
Latino Students			
Boys	63	90	153
Girls	82	97	179
Non-Target Students			
Boys	62	53	115
Girls	59	70	129
Total	528	472	1000

Note. $n_{boys} = 459$.
$n_{girls} = 541$.

[a]Ten years old or younger.

[b]Eleven years old or older.

Experimenters

In each city, visitors were drawn from local universities or school systems and were identified by contacts based on their qualifications and experience working with children. Whenever possible, visitors were placed in classrooms so

that the visitor was from the same ethnic group as the majority of students in that class. However, this was possible in only a few cases because most of the classes participating in the study were integrated classes.

In each city, testers were trained to administer the pre- and posttests. Two of the testers were evaluation project staff members; the other eight testers were graduate students drawn from local universities in each city.

Instruments

Two types of instruments were developed for this study: instruments which were used as pretest/posttest measures and instruments which were used weekly. Some instruments were completed by the students, some by the teacher, and some by the visitor. Those instruments which were completed by the visitor provided information on students' responses to the programs as well as on the amount and kinds of related classroom activities which took place each week.

Student instruments measured:

1. Appeal of the program;
2. Comprehension of the story line;
3. Knowledge of math content;
4. Attitudes toward television, math, and television programs on math; and
5. A range of social attitudes related to series goals.

Teacher instruments measured general attitudes and reactions to each week's program.

Visitors monitored students' attention during the program viewing; they also observed and recorded the

behaviors of students and teachers before and after program viewing, and collected other information about what had happened at other times during the week.

All instruments which were employed in the evaluation were pilot-tested with small groups of target audience students.

Student Attitude Form (Pretest/Posttest). The first side of the Student Attitude Form examined attitudes toward television, math, and television programs on math (see Appendix B). The tester and students read through each item together and students checked boxes if they thought each of the concepts was exciting or boring, fun or no fun, easy or hard, good or bad, and if they didn't like it or liked it.

The second side of the Student Attitude Form measured a range of attitudes related to the cultural and social goals of the series. These included attitudes toward sex roles, adult-child relationships, self-concept, attitudes toward math, attitudes toward one's own culture, attitudes toward other cultures, and attitudes toward television. The tester and the students together read through 15 statements; if a student agreed with the statement, he or she circled yes; if a student disagreed with the statement, he or she circled no.

Student Math Content Form (Pretest/Posttest). The first side of the Student Math Content Form included fourteen statements with which the student either agreed (by circling yes) or disagreed (by circling no) (see Appendix B). These items included statements about graphing, measurement of time, rounding off, weight, scaling, estimation, metric measurement, and mapping.

The second side of the Student Math Content Form included six items which asked for free responses to questions on rounding off and metric weight. Again the tester and students read through the entire instrument together.

Teacher Semantic Differential (Pretest/Posttest). This instrument examined teachers' attitudes toward educational television, math, and television programs on math (see Appendix B). Teachers rated each of these three concepts on twenty pairs of bipolar adjectives, rating them on a seven-point scale.

Teacher Opinion Form (Posttest only). After completing the Teacher Semantic Differential posttest, teachers were asked to complete the Teacher Opinion Form (see Appendix B). Teachers responded to open-ended questions about their opinions on:

1. their overall impressions of the "Infinity Factory" series;
2. those particular parts of the programs which were most effective in meeting the objectives of the series;
3. those parts of the programs which were least effective; and
4. ways in which the television programs and program guides could be improved.

Student Weekly Response Form. Each week, immediately following viewing of "Infinity Factory," students completed a Student Weekly Response Form (see Appendix C). The first side of this instrument measured student appeal for the show overall and for the "Brownstone" segments, "City Flats," "Scoops' Place," and animation. The visitor and students read through the entire instrument together and students checked boxes if they

thought a specific segment was hard or easy, good or no good, fun or no fun, too long or too short, and too fast or too slow. It was determined in the posttesting session that the items "too long or too short," and "too fast or too slow" were misunderstood by about half of the sample, and therefore these items were dropped from the analysis.

The second side of the Student Weekly Response Form measured students' comprehension of story line, knowledge of math content, attitudes toward math, and social attitudes. Ten statements which sampled the objectives of the program in the above areas were used each week. The visitor and students read through the ten statements together. If a student agreed with a statement, he or she circled yes; if a student disagreed with a statement, he or she circled no.

Preliminary examination of the data revealed a positive response bias in students' responses to these items. The mean score on positively worded items was consistently higher than the mean score on negatively worded items. Comparison of mean scores on positive items with mean scores on negative items revealed that this effect was operating equally across all groups. In order to control for this bias, items were not analyzed separately but were aggregated into subscales with equal numbers of positively and negatively worded items.

In addition, two free response items were included when appropriate. Students were able to complete the Student Weekly Response Form in approximately 15 minutes.

Student Attention Form. Visitors were trained to observe and record viewing behavior of students while they watched the television program, using the Student Attention Form (see Appendix C). The visitor observed two groups of five students each, alternating between groups every 15

seconds. The visitor recorded the number of students in each group who were either attending to or actively responding to the program during that 15-second interval. Attending was operationally defined as maintaining eye contact with the television screen; responding was defined as being actively involved with the program, e.g., commenting on the program, laughing, or moving to the music. Audio cues on the Student Attention Form indicated the end of each 15-second segment.

Teacher Weekly Questionnaire Each week the teacher completed a Teacher Weekly Questionnaire (see Appendix C). This instrument was the same for each week and contained 39 statements which teachers checked if they agreed with the statement. Statements dealt with 11 areas related to the program:

1. educational effectiveness;
2. class preparation;
3. program guides;
4. program presentation;
5. use of language;
6. technical quality;
7. student attention;
8. program appeal;
9. math content;
10. math attitudes; and
11. social attitudes.

In addition, teachers were encouraged to write any comments they might have on specific aspects of the program and/or changes they would suggest. Teachers were also asked to report any follow-up activities which took place during the week, and to describe any students' reactions to previous weeks' programs.

Training Procedures

Visitors attended a one-day training session where the objectives of the evaluation study, procedures to be followed, and instruments to be used were explained. Detailed, written instructions for using each instrument were provided, including protocols for instructing students in the use of the Student Weekly Response Forms and for any additional verbal information to be provided students. Videotapes of groups of children watching a television program were used to train visitors in the use of the Student Attention Form and to maximize inter-rater reliability among visitors.

Teachers attended a half-day orientation session where the objectives and scope of the "Infinity Factory" series and the procedures of the evaluation study were explained. Teachers were provided copies of Program Guides for the eight programs (see Appendix M), and were told to use the programs as they saw fit with their students.

Testers were provided with specific instructions on procedures to be followed during the pre- and post-testing sessions.

Experimental Procedures

The programs were broadcast over WGBH-TV for Boston and Lawrence classes, and over KCET-TV for Los Angeles classes. Classes in New York viewed the program using in-school videotape equipment.

Pretest Session. The first week of the evaluation, testers arrived at the classroom at specified times which were convenient for the class. The tester gave the teachers the Teacher Semantic Differential to be filled

out while the students completed the Student Attitude Form and then the Student Math Content Form. Each instrument took approximately 15 minutes. When the tester distributed the Student Attitude Form, he or she asked students to put their first name and last initial at the top of the page. Then the students were asked to give their true feelings in a number of different areas, and told that there were no right or wrong answers to the statements to be read. Students were assured that their teachers, principals, and parents would not see their responses. When the Student Math Content Form was distributed to students, they were instructed to leave a statement unanswered if they did not know the answer.

Experimental Sessions. For each of the eight weekly viewing sessions, the classroom visitor arrived in the classroom about 15 minutes before the program was scheduled to begin. The visitor recorded any related classroom activities which occurred before the program.

The class viewed the half-hour television program either in their classroom or in another room in the school. During the viewing session, the visitor recorded eye contact and active responses for two groups of five students each, alternating from one group to another at 15-second intervals, using the Student Attention Form.

Immediately following program viewing, the visitor and students completed the Student Weekly Response Form. The visitor and students read through all items together as the students filled out this form. At the same time, the teacher completed a Teacher Weekly Questionnaire.

Afterwards, the visitor observed and recorded any follow-up classroom activities. Before leaving the classroom, the visitor reminded the teacher to take note of

any related activities during the coming week, and to take note of any incidental references to the program made by the students.

Visitors monitored what happened in the classroom before, during, and after program viewing. They reported and categorized any related activities which occurred before or after the program, noting whether the teacher dealt with the math content, cultural, or other attitudinal areas of the program and whether the teacher used any suggestions from the Program Guides.

Posttest Session. During the week following broadcast of the eighth program, the tester who had administered the pretest returned to the same classroom and repeated the pretest procedures for the Student Attitude Form, Student Math Content Form, and Teacher Semantic Differential. At this time the teacher also completed the Teacher Opinion Form.

Limitations of the Study

Several factors have been identified which limit the generalizability of this study:

1. Technical problems arose in all cities. In some cases, stations broadcast programs at the wrong time or in incorrect sequence. Problems with both hardware and software came up in classes which used portable videotape equipment to view the program.
2. The sample for the study was selected by school administrators on very short notice. It is unlikely that the sample selection procedure was exactly random, but there is reasonable con-

fidence that the sample is sufficiently representative of the population.

3. Background data were collected only on students' age, sex, and ethnic group.
4. The forced choice format used for student instruments has the advantages of being easy for students to understand and controlling for reading ability. However, such a format gives less detailed information than free-response formats about attitudes and understanding of math content.
5. Although instruments and procedures were designed so that immediate and individual responses of students could be obtained, visitors reported that in some classroom settings minor problems were encountered with interference from peers or teachers.
6. It was recognized that a potential difficulty in measuring appeal is the tendency for students to report what they think adults would like to hear.
7. To compensate for a detected positive response bias, students' responses were aggregated into subscales with equal numbers of positive and negative items.
8. Test items had content and face validity. Time constraints did not allow for extensive, rigorous test construction procedures (e.g., item analysis), resulting in minor problems with some test items which were later accounted for in the data analysis.
9. In a 10-week study, problems were expected with mortality. Out of a total sample of over 1,000 students, the mean number viewing each week was

800. The sample for pretest/posttest analysis consisted of the 660 students who completed pretest/posttest measures and saw at least six of the eight programs. The sample for analysis of comprehension, math content, and attitude subscales consisted of the 383 students who saw all eight programs. The sample for analysis of attention trends consisted of 21 classes for which data from all eight programs were available, and the sample for trend analysis of teacher responses consisted of 27 teachers who completed Teacher Weekly Questionnaries for all eight programs.

10. In some cities, the program was available for home viewing during some weeks. Although students were not informed of this, a number reported that they had watched some programs at home.

3. ANALYSIS OF THE DATA

Data were analyzed using computer programs from the Statistical Package for the Social Sciences (SPSS). Programs from these packages provided frequency data, statistical analyses, and, in some cases, posthoc analyses. The value of $p < .05$ was selected as a minimum level of statistical significance.

Student Attention

An attention score was calculated from the mean percentage of children attending to and/or actively responding to each 15-second segment. From this score the mean percentage of attention was calculated for each show overall and for each major segment.

Mean percentages of attention for the eight shows were compared using a fixed effects, 21 (classes) X 8 (shows), repeated measures analysis of variance. The two-way model was used in order to examine differences in overall attention among shows for the entire sample, after any differences which could be attributed to differences among classes was removed. Differences among classes were expected; they were considered to be the result of both actual difference among groups and differences in the way visitors recorded attention, and were reported but not discussed.

The sample for this analysis consisted of the 21 classes for which attention data were available for all eight shows. Posthoc contrasts using the Scheffé test were performed to compare attention among the eight shows.

Student Appeal

Student appeal was determined from responses to the first side of the Student Weekly Response Forms (see Appendix C). Students rated five areas of each sho (the show overall, "Brownstone" segments, "City Flats," "Scoops' Place," and animation segments) as being hard or easy, good or no good, and fun or no fun. Responses on the adjective pairs were aggregated into subscales;

1. Weekly Subscales for the five areas of each show, based on the number of positive adjectives checked for each particular week (the 40 Weekly Subscales range from 0 to 3);
2. Total Weekly Subscales, based on the total number of positive adjectives checked for all five areas for a particular week (the eight Total Weekly Subscales ranged from 0 to 15);
3. Total Series Subscales for the five areas, based on the total number of positive adjectives checked for the total of eight weeks (the five Total Series Subscales ranged from 0 to 24).

Comparisons of shows were made by analyzing the Weekly Subscales and the Total Weekly Subscale mean scores using one-way analyses of variance. A test for linear trend was made where appropriate.

Comparisons of groups were made by analyzing the five Total Series Subscales using a 3 (ethnic group) X 2 (age) X 2 (sex) analysis of variance.

Where significant F ratios were found, posthoc comparisons were performed using the Scheffé test.

Comprehension of Dramatic Story Line

Students' comprehension of the story line was assessed by analysis of a Comprehension Subscale constructed from students' responses to comprehension items on the

Student Weekly Response Forms.* In order to control for positive response bias, 12 positively worded and 12 negatively worded items were randomly selected for the Comprehension Subscale (see Appendix D).

Comprehension subscale scores were analyzed using a fixed effects, 3 (ethnic groups) X 2 (age) X 2 (sex) analysis of variance. Where there were significant ethnic group effects, posthoc contrasts were performed using the Scheffé test on mean subscale scores.

Knowledge of Math Content

Effectiveness of the eight-show treatment in improving students' knowledge of math content was assessed in two ways:

1. by comparison of the pretest and posttest results on the Student Math Content Form; and
2. by an analysis of a math content subscale constructed from the students' responses on the eight Student Weekly Response Forms.

The sample for the pretest/posttest analysis consisted of all students who had completed both the pretest and the posttest, and who had seen at least six of the eight shows. The entry levels of the various groups in the sample were determined by analyzing mean pretest scores using a fixed effects, 3 (ethnic group) X 2 (age) X 2 (sex) analysis of variance.

Pretest and posttest scores on the Math Content Form were compared by performing a t-test on the overall mean gain score (posttest score - pretest score).

Then an analysis of gain scores was performed, controlling for students' performances on the Math Content Form pretest. Content gain scores were

*Items from the Student Weekly Response Form, Show C, were dropped from the analysis of Comprehension, Math Content, and Attitude Subscales because 31% of the classes had reported audio or video problems for that show.

analyzed using a 3 (ethnic group) X 2 (age) X 2 (sex) analysis of covariance, with pretest scores as the covariate. Any variation in gain scores attributable to pretest score differences was removed before analyzing the effects of ethnic group, age, sex, and interaction terms.

Where there were significant ethnic group effects, posthoc contrasts were performed using the Scheffé test. For the analysis of covariance, this analysis was performed on the adjusted means.

The Math Content Subscale was constructed from items on math content on the eight Student Weekly Response Forms. In order to control for positive response bias, four positively worded items, four negatively worded items, and four free response items were randomly selected for this subscale (see Appendix D).

Analysis of the Math Content Subscale was carried out in the same way as the analysis of the Comprehension Subscale reported in the previous section.

Student Attitudes

Effectiveness of the eight-show treatment in influencing students' attitudes was assessed in two ways:

1. by comparison of the pretest and posttest results on the Student Attitude Form; and
2. by an analysis of an Attitude Subscale constructed from students' responses on the eight Student Weekly Response Forms.

Analysis of pretest/posttest differences on the Student Attitude Form was carried out in the same way as the analysis of the Student Math Content Form, described in the previous section.

The Attitude Subscale was constructed from items on social attitudes and attitudes toward math on the eight

Student Weekly Response Forms. In order to control for positive response bias, nine positively and nine negatively worded items were randomly selected for this subscale (see Appendix D).

Analysis of student responses on the 18-item Attitude Subscale was carried out in the same way as the analysis of student responses on the Comprehension subscale described above.

Teacher Responses

Differences in teachers' attitudes toward educational television, math, and television programs on math were assessed by comparison of pretest and posttest differences on the Teacher Semantic Differential. Each of the three concepts was rated for 20 pairs of bipolar adjectives on a scale of one (least positive) to seven (most positive). The mean score for all teachers for the 20 adjective pairs was computed for each concept. A two-tailed, correlated t-test was used to compare differences on the pretest and posttest scores.

Teachers' attitudes toward the eight "Infinity Factory" programs were assessed by a comparison of teachers' responses to the 39 items on the Teacher Weekly Questionnaire (see Appendix C). Ten subscales were constructed from 37 items:

1. Class Preparation;
2. Program Guides;
3. Program Presentation;
4. Language; Use of Language;
5. Technical Quality;
6. Student Attention;
7. Student Appeal;
8. Math Content;

9. Math Attitudes;

10. Social Attitudes.

The items which made up each subscale are listed in Appendix E.

Item 37 ("This program is educationally effective"), and Item 38 ("The overall presentation in this week's program was outstanding (=4), good (=3), mediocre (=2), or poor (=1)") were analyzed separately.

Teacher subscales and items 37 and 38 were compared using a fixed effects, 8 (show) X 27 (teacher), repeated measures analysis of variance, with a test for linear trends over shows. The two-way model was used in order to examine differences among shows after any differences which could be attributed to teachers were removed. Differences among teachers were expected and were reported but not discussed. The sample for this analysis included all teachers who had completed Teacher Weekly Questionnaires for all eight shows. Where significant F ratios were found for the main effect of show, posthoc contrasts were performed using the Scheffé test.

Teachers' responses to the four questions on the Teacher Opinion Form were categorized, tabulated, and analyzed informally.

4. RESULTS

Student Attention

Mean percentages of student attention over eight shows for the 21 classes for which complete attention data were available are shown in Appendix F, Table F.1.

Two-way analysis of variance (see Table 4.1) revealed significant differences in mean attention among classes ($p < .001$) and a significant difference over eight shows ($p < .05$). There was no significant linear trend over eight shows at the .05 level (see Figure 1).

Although the main effect of difference among shows is significant, Scheffé posthoc analyses revealed that when mean percentages of attention are compared, no difference between any two shows is significant at the .05 level.

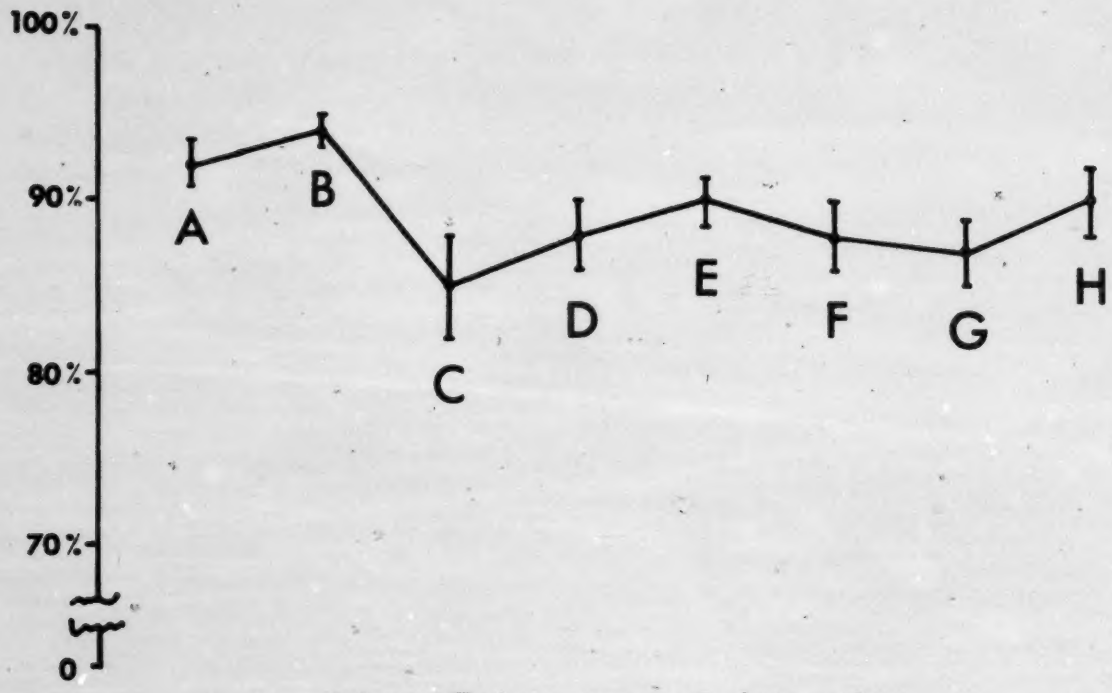
Table 4.1

Student Attention Trends: Analysis of Variance

Source of Variation	SS	df	MS	F
Main Effects	17.76	27	.65	4.32***
Class	15.02	20	.75	4.93***
Show	2.74	7	.39	2.57*
Linear Term	.37	1	.37	2.46
Deviation from Linear	2.37	6	.39	2.60
Explained	17.76	27	.65	4.32***
Error	21.30	140	.15	
Total	39.07	167	.23	

*$p < .05$.

***$p < .001$.



Mean Percentage of Attention
Shows A to H

Figure 4.1

Student Appeal

Comparisons of Shows. Reported below are five one-way analyses of variance performed on the Weekly Appeal Subscales for the show overall, "Brownstone" segments, "City Flats," "Scoops' Place," and animation segments. Also reported are the results of tests for linear trend over shows and Scheffé posthoc analyses on the five Weekly Subscales. Descriptive statistics and results of five one-way analyses of variance are shown in Appendix G, Tables G.1 to G.10. These results are summarized in Table 4.2 below, which presents F ratios, degrees of freedom, and levels of significance for show effect and linear trend over shows.

Table 4.2

Comparisons of Student Appeal Among Shows:

Summary of Results

Weekly Appeal Subscale	Show		Linear Trend	
	df	F	df	F
Show Overall	(7, 6474)	21.10***	(1, 6)	44.63***
Brownstone	(7, 6369)	16.68***	(1, 6)	97.62***
City Flats	(7, 6398)	3.31**	(1, 6)	4.51*
Scoops' Place	(7, 6362)	15.20***	(1, 6)	5.08*
Animation	(6, 5519)	6.23***	(1, 5)	15.20***

*$p < .05$

**$p < .01$

***$p < .001$.

The five one-way analyses of variance revealed significant differences among shows and significant linear trends on all five Weekly Appeal Subscales. Scheffé post-hoc analyses revealed that when Weekly Appeal Subscale scores are compared between shows pairwise, the following differences are significant ($p < .05$):

1. Mean Weekly Appeal Subscale scores for the show overall for Shows A (2.84) and B (2.80) were significantly higher than for the other shows, and mean show overall Weekly Subscale scores for Shows G (2.55) and H (2.59) were significantly lower than for the other shows;
2. Mean "Brownstone" Weekly Appeal Subscale scores were significanly higher for Shows A (2.75) and

B (2.71) than for the other shows, and significantly lower for Shows G (2.46) and H (2.49) than for the other shows;

3. Mean "City Flats" Weekly Appeal Subscale score for Show A (2.70) was significantly higher than Show G (2.58);
4. Mean "Scoops' Place" Weekly Appeal Subscale score was significantly higher for Show C (2.71) than Show A (2.37);
5. Mean animation Weekly Subscale score was significantly lower for Show G (2.60) than for Show D (2.72), Show B (2.74), and Show F (2.77); and significantly lower for Show H (2.63) than Show F.

The Total Weekly Appeal Subscale was computed from the sum of the five Weekly Appeal Subscales analyzed above. Descriptive statistics and results of one-way analysis of variance, including test for linear trend over shows, performed on the Total Weekly Appeal Subscale, are shown in Appendix G, Tables G.11 and G.12.

One-way analysis of variance on the Total Weekly Subscale revealed a significant difference among shows, $F(7, 5794) = 6.69$, $p < .001$. There was a significant downward linear trend over shows, $F(1, 6) = 18.35$, $p < .001$. Scheffé posthoc analyses revealed that when Total Weekly Subscale scores are compared pairwise, Show G (mean = 12.93) is rated significantly lower than Shows C, D, F, or B (means = 13.46, 13.50, 13.58, and 13.67, respectively); Show A (mean = 13.35) is rated significantly lower than Show B.

Comparisons of Groups. Reported below are the results of five three-way analyses of variance (3(ethnic group) X 2(age) X 2(sex)) performed on the Total Series

Appeal Subscales for the show overall, "Brownstone" segments, "City Flats," "Scoops' Place," and animation segments. Descriptive statistics and results of the five three-way analyses of variance are shown in Appendix H, Tables H.1 to H.10.

Three-way analysis of variance on Total Series Appeal Subscale scores for the shows overall revealed a significant difference among mean scores for Black (21.92), Latino (21.53), and non-target (20.34) students, $F(2, 273) = 4.72$, $p < .01$. Scheffé posthoc analyses revealed significant differences between Black and non-target students' Total Series Subscale scores, and between Latino and non-target students' scores. However, the difference between Black and Latino students was not significant at the .05 level.

There was no significant difference between younger and older students' scores on Total Series Subscale for the shows overall (means = 21.60 and 20.97, respectively), $F(1, 273) = 12.21$. $p < .05$. However, the Total Series Subscale score was significantly higher for girls (mean = 21.96) than for boys (mean = 20.56), $F(1, 273) = 126.91$, $p < .001$.

There was a significant interaction between ethnic group and sex, $F(2, 273) = 74.98$, $p < .001$. The Total Series Subscale scores for the shows overall were higher for Latino (mean = 22.00) and non-target (mean = 21.96) girls than for Latino (mean = 20.86) and non-target (mean = 18.78) boys. However, mean Total Series Subscale scores were slightly lower for Black girls (21.89) than for Black boys (21.99). There were no other significant two-way interactions on the Total Series Subscale for the shows overall.

Three-way analysis of variance on "Brownstone" Total Series Appeal Subscale scores revealed significant differences

among mean scores for Black (21.89), Latino (21.33), and non-target (19.60) students, $F(2, 252) = 95.78$, $p < .001$. Scheffé posthoc analyses revealed a significant difference between Black and non-target students' scores ($p < .05$). There were no significant differences between Total Series Subscale scores for Black and Latino students or between Latino and non-target students.

There was no significant difference between younger and older students' scores (means = 21.49 and 20.40, respectively) on the "Brownstone" Total Series Subscale, $F(1, 252) = 2.96$, $p > .05$. Mean subscale scores were significantly higher for girls (21.88) than boys (19.99), $F(1, 252) = 16.83$, $p < .001$.

There was a significant interaction between ethnic group and sex, $F(2, 252) = 6.20$, $p < .01$. Total Series Subscale scores were higher for Latino and non-target girls (means = 22.28 and 21.44, respectively) than for Latino and non-target boys (means = 20.27 and 17.60, respectively). However, Total Series Subscale scores were slightly lower for Black girls (mean = 21.76) than for Black boys (mean = 21.90).

Three-way analyses of variance on "City Flats" Total Series Subscale scores revealed significant differences among ethnic groups, $F(2, 248) = 3.44$, $p < .01$. Scheffé posthoc analyses revealed that the mean Total Series Subscale scores were significantly lower for non-target students (20.34) than for Black (21.52) and Latino (21.52) students, $F(1, 183) = 6.18$, $p < .05$.

There was no significant difference between "City Flats" Total Series Subscale scores for younger and older students (means = 21.52 and 21.79, respectively), $F(1, 248) = .14$, $p > .05$. Mean subscale scores were significantly higher for girls (21.54) than for boys (20.71), $F(1, 248) = 3.95$, $p < .05$.

There was a significant interaction between ethnic group and age, $F(2, 248) = 4.08$, $p < .05$. Older Black and Latino students (means = 22.12 and 21.59, respectively) rated "City Flats" higher than younger Black and Latino students (means = 21.20 and 21.38) respectively). However, younger non-target students (mean = 21.26) rated "City Flats" higher than older non-target students (mean = 19.54).

There was also a significant interaction between ethnic group and sex $F(2. 248) = 2.20$, $p < .05$. "City Flats" Total Series Subscale scores were higher for Latino and non-target girls (means = 21.92 and 21.24, respectively) than for Latino and non-target boys (means = 20.94 and 19.42, respectively). However, Total Series Subscale scores were slightly lower for Black girls (mean = 21.44) than for Black boys (mean = 21.56).

There was no significant interaction between age and sex on the "City Flats" Total Series Subscale score.

Three-way analysis of variance on "Scoops' Place" Total Series Subscale scores revealed significant differences among ethnic groups. $F(2, 238) = 6.82$, $p < .001$. Scheffé posthoc analyses revealed that Total Series Subscale scores for Black and Latino students (means = 21,41 and 21.06, respectively) were not significantly different, but that scores for both groups were significantly higher than for non-target students (mean = 19.52).

There was no significant difference between "Scoops' Place" Total Series Subscale scores for younger and older students (means = 20.77 and 20.70, respectively), $F(1, 238) = .08$, $p > .05$. There was no significant difference between "Scoops' Place" subscale scores for boys and girls (means = 20.28 and 21.13, respectively), $F(1, 238) = 3.32$, $p > .05$.

There were no significant two-way interactions on "Scoops' Place" Total Series Subscale scores.

The Total Series Subscale for animation segments was calculated from the sum of weekly appeal subscales for Show B to Show H. (Responses for Show A were dropped from analysis because there was no cartoon animation in Show A.) The score on the 21 items from Shows B through H was then multiplied by 8/7 to achieve a score with a maximum of 24 in order to facilitate comparisons with other Total Series Subscales.

The grand mean for the animation Total Series Subscale was 21.65. There were no significant differences among ethnic groups. $F(2, 270) = 1.55$, $p > .05$; ages, $F(1, 270) = .92$, $p > .05$; and sexes, $F(1, 270) = 1.77$, $p > .05$. There were no significant two-way interactions among variables.

Comprehension of Dramatic Story Line

Frequency data for the Comprehension Subscale are shown in Appendix I, Tables I.1. The grand mean for all subgroups was 77.7.*

Three-way analysis of variance (see Table 4.3) revealed no significant differences ($p > .05$) among Comprehension Subscale scores for Black (76.6), Latino (78.4), and non-target (78.7) students. Mean Comprehension Subscale scores were significantly higher ($p < .01$) for older (80.08) than for younger (75.99) students. There was no significant difference ($p > .05$) between mean Comprehension Subscale scores for girls (78.24) and boys (77.1).

There were no significant two-way interactions on Comprehension Subscale scores.

---

*Comprehension, Math Content, and Attitude Subscale scores were converted to a 100-point scale to facilitate comparisons among subscales.

Table 4.3

Analysis of Variance: Comprehension Subscale

Source of Variation	SS	df	MS	F
Main Effects	1616.14	4	404.03	3.08*
Ethnic Group	113.38	2	56.69	0.43
Age	1192.89	1	1192.89	9.09**
Sex	89.91	1	89.91	0.68
Two-Way Interactions	852.45	5	170.49	1.30
Ethnic Group X Age	199.64	2	99.82	0.76
Ethnic Group X Sex	606.99	2	303.49	2.31
Age X Sex	21.56	1	21.56	0.16
Explained	2468.60	9	274.28	2.09*
Error	43924.91	335	131.11	
Total	46393.52	344	134.86	

Note. n = 345 = all students for whom complete data were available.

*p < .05

**p < .01

Knowledge of Math Content

Pretest/Posttest Comparisons. Descriptive statistics for math content pretest and gain scores are shown in Appendix J, Tables J.1 and J.2.

Three-way analysis of variance on math content pretest scores (see Table 4.4 below) revealed no significant differences ($p > .05$) among pretest scores for Black, Latino, and non-target students (means = 9.41, 9.13, and 9.38, respectively). Pretest scores were significantly greater ($p < .001$) for older (mean = 9.92) than younger (mean = 8.70) students; and significantly greater ($p < .001$) for boys (mean = 9.71) than girls (mean = 9.02).

Table 4.4

Analysis of Variance: Math Content Pretest

Source of Variation	SS	df	MS	F
Main Effects	340.56	4	85.14	11.46***
Ethnic Group	31.16	2	15.58	2.09
Age	253.85	1	253.85	34.18***
Sex	83.34	1	83.34	11.22***
Two-Way Interactions	68.87	5	13.77	1.85
Ethnic Group X Age	16.21	2	8.10	1.09
Ethnic Group X Sex	50.04	2	25.02	3.36*
Age X Sex	5.49	1	5.49	0.74
Explained	409.44	9	45.49	6.12***
Error	4827.52	650	7.42	
Total	5236.96	659	7.94	

Note. n = 660 = all students who completed pretest and posttest and saw at least six of the eight shows.

*$p < .05$

***$p < .001$

There was a significant interaction ($p < .05$) between ethnic group and sex. While pretest scores were higher for boys than girls for all ethnic groups, the difference between non-target boys and girls was considerably greater than between Black and Latino boys and girls. Figure 4.2 illustrates this interaction. No other two-way interactions were significant at the .05 level.

MATH CONTENT MEAN PRETEST SCORES:
INTERACTION OF ETHNIC GROUP AND SEX

Figure 4.2

The overall mean gain score (3.07) was significant at the $p < .001$ level, t (659) = 26.10.

Three-way analysis of covariance (see Table 4.5 below) revealed a significant difference ($p < .001$) among adjusted gain scores for Black (2.73), Latino (2.98), and non-target (3.75) students. Scheffé posthoc analyses revealed that the differences between adjusted mean gain scores for Black and Latino students was not significant, and that non-target students' adjusted mean gain scores were significantly greater ($p < .05$) than Black and Latino students' adjusted mean gain scores.

Adjusted mean gain scores were significantly higher ($p < .001$) for older (3.50) than for younger (2.68) students. The difference between adjusted mean gain scores for boys (3.25) and girls (2.94) was not significant ($p > .05$).

Table 4.5
Analysis of Covariance: Math Content Gain Scores

Source of Variation	SS	df	MS	F
Covariate				
Content Pretest Score	1896.08	1	1896.08	331.69***
Main Effects	250.49	4	62.62	10.95***
Ethnic Group	108.82	2	54.41	9.51***
Age	101.93	1	101.93	17.83***
Sex	15.32	1	15.32	2.68
Two-Way Interactions	170.41	5	34.08	5.96***
Ethnic Group X Age	4.97	2	2.48	0.43
Ethnic Group X Sex	158.32	2	79.16	13.84***
Age X Sex	17.93	1	17.93	3.13
Explained	2316.99	10	231.69	40.53***
Error	3709.87	649	5.71	
Total	6026.87	659	9.14	

***$p < .001$

There was a significant two-way interaction between ethnic group and sex ($p < .001$). Gain scores were higher for Black and non-target girls (means = 2.78 and 4.54, respectively) than for boys (means = 2.30 and 2.90, respectively). However, mean gain scores were higher for Latino boys (4.01) than girls (2.58). There were no other significant two-way interactions.

Math Content Subscale. Descriptive statistics for the Math Content Subscale are shown in Appendix J, Table J.3.

Three-way analysis of variance (see Table 4.6) revealed a significant difference ($p < .001$) among mean subscale scores for Black, Latino, and non-target students

(means = 71.2, 71.8, and 80.1, respectively). Scheffé posthoc analyses revealed that the difference between Math Content Subscale scores for Black and Latino students was not significant ($p > .05$); however, differences between Black and non-target students, and between Latino and non-target students were significant ($p < .05$).

Table 4.6

Analysis of Variance: Math Content Subscale

Source of Variation	SS	df	MS	F
Main Effects	15112.00	4	3778.00	16.49***
Ethnic Group	4043.65	2	2021.82	8.82***
Age	9186.46	1	9186.46	40.11***
Sex	23.27	1	23.27	0.10
Two-Way Interactions	275.93	5	55.18	0.24
Ethnic Group X Age	186.85	2	93.43	0.40
Ethnic Group X Sex	95.84	2	47.92	0.20
Age X Sex	0.82	1	0.82	0.00
Explained	15387.93	9	1709.77	7.46***
Error	85415.81	373	228.99	
Total	100803.75	382	263.88	

Note. $n = 383$ = all students for whom complete data were available.

***$p < .001$

Math Content Subscale scores were significantly higher ($p < .001$) for older (mean = 79.8) than younger (mean = 69.0) students. There was no significant difference ($p > .05$) between boys' and girls' subscale scores (means = 73.5 and 74.1, respectively). There were no significant two-way interactions ($p > .05$).

Student Attitudes

Pretest/Posttest Comparisons. Descriptive statistics for pretest and gain scores on four areas of student attitudes (social and math attitudes related to series goals, attitudes toward television, attitudes toward math, and attitudes toward television programs on math) are shown in Appendix K, Tables K.1 to K.8.

The overall mean gain score (.63) on the 14-item measure of social and math attitudes related to program goals was significant at the $p < .001$ level, $t(659) = 8.12$. Three-way analysis of covariance on mean gain scores (see Table 4.7 below) revealed significant differences ($p < .05$) among adjusted mean gain scores for Black (.83), Latino (.57), and non-target (.38), students. Scheffé posthoc analyses revealed that the difference between adjusted mean gain scores for Black and non-target students was significant ($p < .05$). However, differences between adjusted mean gain scores for Black and Latino students, and for Latino and non-target students, were not significant at the .05 level.

Adjusted mean gain scores for social and math attitudes were significantly higher for girls (1.01) than boys (.12). There were no significant differences between older and younger students' gain scores (adjusted means = .50 and .75, respectively). There were no significant two-way interactions on social and math attitudes gain scores.

Table 4.7

Analysis of Covariance: Gain Scores, Social and Math Attitudes

Source of Variation	SS	df	MS	F
Covariate				
Pretest Score	563.96	1	563.96	194.26***
Main Effects	146.37	4	36.59	12.60***
Ethnic Group	22.60	2	11.30	3.89*
Age	10.38	1	10.38	3.58
Sex	102.59	1	102.59	35.34***
Two-Way Interactions	19.02	5	3.80	1.31
Ethnic Group X Age	8.00	2	4.00	1.38
Ethnic Group X Sex	2.75	2	1.37	0.47
Age X Sex	8.04	1	8.04	2.77
Explained	729.35	10	72.94	18.40***
Error	1882.63	649	2.90	
Total	2611.98	659	3.96	

*$p < .05$
**$p < .01$
***$p < .001$

There was no significant difference between overall mean pretest and posttest scores on attitudes toward television, $t(606) = .50$, $p > .05$. In addition, three-way analysis of covariance on these mean gain scores (see Table 4.8 below) revealed no significant differences among ethnic groups, age groups, and sexes; and no significant two-way interactions between these main effects ($p > .05$ for all main effects and interactions).

Table 4.8

Analysis of Covariance: Gain Scores, Attitudes Toward Television

Source of Variation	SS	df	MS	
Covariate				
Pretest Score	157.62	1	157.62	297.82***
Main Effects	2.01	4	0.50	0.94
Ethnic Group	0.76	2	0.38	0.71
Age	0.10	1	0.10	0.18
Sex	1.10	1	1.10	2.07
Two-Way Interactions	1.60	5	0.32	0.60
Ethnic Group X Age	0.36	2	0.18	0.33
Ethnic Group X Sex	0.02	2	0.01	0.01
Age X Sex	1.26	1	1.26	2.37
Explained	161.22	10	16.12	30.25***
Error	317.54	596	0.53	
Total	478.76	606	0.79	

*$p < .05$
**$p < .01$
***$p < .001$

There was no significant difference between overall mean pretest and posttest scores on attitudes toward math, $t(589) = .53$, $p > .05$. In addition, three-way analysis of covariance on these mean gain scores (see Table 4.9 below) revealed no significant differences among ethnic groups, age groups, and sexes; and no significant two-way interactions between these main effects ($p > .05$ for all main effects and interactions).

Table 4.9

Analysis of Covariance
Gain Scores, Student Attitudes Toward Math

Source of Variation	SS	df	MS	F
Covariate				
Pretest Score	277.27	1	277.27	111.66***
Main Effects	12.22	4	3.06	1.23
Ethnic Group	6.60	2	3.30	1.33
Age	2.19	1	2.19	0.88
Sex	3.95	1	3.95	1.59
Two-Way Interactions	5.72	5	1.14	0.46
Ethnic Group X Age	0.28	2	0.14	0.05
Ethnic Group X Sex	5.58	2	2.79	1.12
Age X Sex	0.06	1	0.06	0.02
Explained	295.21	10	29.52	11.90***
Error	1433.91	579	2.48	
Total	1729.12	589	2.94	

*$p < .05$
**$p < .01$
***$p < .001$

The overall mean gain score (.46) on attitudes toward television programs on math was significant at the $p < .001$ level, $t(592) = 5.76$. Three-way analysis of covariance on these mean gain scores (see Table 4.10 below) revealed significant differences ($p < .05$) among adjusted mean gain scores for Black (.67), Latino (.42), and non-target (.19), students. Scheffé posthoc analyses revealed that the difference between adjusted mean gain scores for Black and non-target students was significant ($p < .05$); however, differences between Black and Latino students, and between Latino and non-target students

were not significant ($p > .05$). There were no significant differences between age groups or between boys and girls ($p > .05$).

There was a significant ($p < .01$), two-way interaction between ethnic group and age. Mean gain scores for attitudes toward television programs on math were higher for older Black and Latino students (means = .91 and .33, respectively) than for younger Black and Latino students (means = .45 and .25, respectively). However, mean gain scores were higher for younger (.50) than older (.36) non-target students.

Table 4.10

Analysis of Covariance
Gain Scores, Student Attitudes Toward TV Programs on Math

Source of Variation	SS	df	MS	F
Covariate				
Pretest Score	774.27	1	774.27	318.00***
Main Effects	27.61	4	6.90	2.84*
Ethnic Group	21.12	2	10.56	4.34*
Age	0.97	1	0.97	0.40
Sex	4.89	1	4.89	2.01
Two-Way Interactions	37.40	5	7.48	3.07**
Ethnic Group X Age	28.30	2	14.15	5.81**
Ethnic Group X Sex	11.68	2	5.84	2.40
Age X Sex	0.43	1	0.43	0.18
Explained	839.28	10	83.93	34.54***
Error	1414.05	582	2.43	
Total	2253.33	592	3.81	

*$p < .05$
**$p < .01$
***$p < .001$

Student Attitude Subscale. Descriptive statistics for the Student Attitude Subscale scores are shown in Appendix K, Table K.9.

Three-way analysis of variance (see Table 4.11) revealed significant differences ($p < .05$) among mean Attitude Subscale scores for Black (76.4), Latino (74.2), and non-target (72.40), students. Scheffé posthoc analyses revealed that no difference between mean subscale scores for Black, Latino, and non-target students was significant ($p > .05$).

Mean Attitude Subscale scores were significantly higher ($p < .01$) for older (78.6) than younger (73.2) students. There was no significant difference ($p > .05$) between Attitude Subscale scores for boys (73.1) and girls (75.6). There were no significant two-way interactions ($p > .05$).

Table 4.11

Analysis of Variance: Attitude Subscale

Source of Variation	SS	df	MS	F
Main Effects	2858.48	4	714.62	4.38**
Ethnic Group	1331.74	2	665.87	4.08*
Age	1350.04	1	1350.04	8.26**
Sex	584.13	1	584.13	3.58
Two-Way Interactions	1677.42	5	335.48	2.05
Ethnic Group X Age	705.36	2	352.68	2.16
Ethnic Group X Sex	626.22	2	313.11	1.92
Age X Sex	180.20	1	180.20	1.10
Explained	4535.91	9	503.99	3.09**
Error	51123.91	313	163.34	
Total	55659.82	322	172.86	

Note. $n = 323$ = all students for whom complete data were available.

*$p < .05$
**$p < .01$

Teacher Responses

Pretest/Posttest Comparisons. Pretest and posttest scores on the Teacher Semantic Differential and their correlated *t*-test analyses are shown in Appendix L, Table L.1.

Pretest means on all three concepts (educational television, math, and television programs on math) were above 5.20 on a scale of 1.00 (least positive) to 7.00 (most positive).

While all of the pretest-posttest differences were in the expected positive direction, only the educational television concept score increased significantly from pretest to posttest, $t(33) = 2.72$, p .0.1. There were no significant pretest-posttest differences on concept scores for math, $t(33) = 1.47$, $p > .05$; and television programs on math, $t(33) = 1.37$, $p > .05$.

Of the 36 teachers who completed the open-ended Teacher Opinion Form during posttesting, 31 (86%) described their overall opinion of the series in positive terms. Teachers commented favorably on the following aspects of the series: cultural and social aspects (15 teachers); student appeal and general presentation (12); math content (10); and motivation toward math (8). (Some teachers commented on more than one area.)

Teachers rated the series most effective in relating math to real life (10 teachers); presenting positive social and cultural images (6); and introducing the metric system (4).

Thirteen teachers asked for more emphasis on basic math content. Four teachers suggested that less emphasis be placed on the use of non-standard English.

Particular segments rated effective included: "Scoops' Place" (11 teachers); "City Flats" (10); animation segments (9); "Math in the Street" (5); and "Brownstone" segments (3).

Particular show segments rated ineffective included: "Math Fact" segments (11 teachers); historical segments (9); "Math in the Street" (5); and "Brownstone" segments (3).

Comparison of Shows. Teacher Weekly Questionnaires for all eight weeks were completed by 27 teachers (69% of the total sample). Frequency data and results of two-way, repeated measures analyses of variance (27 (teachers) X 8 shows)) on items 37 and 38 are shown in Appendix L. Tables L.2 to L.5.

The mean score on teachers' rating of the overall presentation (item 38) was 3.0 on a scale of 1 (poor) to 4 (outstanding). There was a significant difference among teachers' responses, $F(26, 173) = 3.42$, $p < .001$. There was also a significant difference among shows, $F(7, 173) = 3.56$, $p < .001$. There was no significant linear trend over shows, $F(1, 6) = .52$, $p > .05$. Scheffé posthoc analyses comparing the means on all shows pairwise revealed that no difference between means for any pair of shows was significant at the .05 level.

The mean rating for educational effectiveness for the eight shows (item 37) was 77%. There was a significant teacher effect, $F(26, 173) = 4.66$, $p < .001$. There were no significant differences among the eight shows, $F(7, 173) = 1.68$, $p > .05$; and no significant linear trend, $F(1, 6) = .04$, $p > .05$.

Frequency data and results of the ten two-way, repeated measures analyses of variance (27 teachers X 8 shows) performed on the ten Teacher Weekly Questionnaire

subscales are shown in Appendix L, Tables L.6 to L.25. These results are summarized in Table 4.12, which presents F ratios, degrees of freedom, and levels of significance for teacher effect, show effect, and linear trend over shows; and overall means for each of the ten subscales.

Table 4.12

Teacher Subscales: Summary of Analyses

Subscale	M	F Ratios		
		Teacher (df=(26,182))	Show (df=(7,182))	Linear Trend (df = (1, 6))
Class Preparation	43%	2.97***	1.08	.65
Program Guide	69%	7.61***	1.58	1.27
Program Presentation	77%	10.42***	2.95**	.40
Language	86%	3.18***	1.20	1.86
Technical Reception	75%	4.11***	2.66*	10.29**
Student Attention	84%	2.10**	3.19***	.44
Program Appeal	81%	7.02***	2.89**	.03
Math Content	75%	2.98***	2.08*	.29
Math Attitudes	82%	3.50***	1.18	.26
Social Attitudes	71%	12.94***	3.00**	.02

*p < .05
**p < .01
***p < .001.

These analyses revealed a significant difference among teachers on all ten subscales, as was expected. There were also significant differences among shows on the

Program Presentation, Technical Reception, Student Attention, Program Appeal, Math Content, and Social Attitudes Subscales. There was a significant linear trend over shows only for the Technical Reception Subscale; this trend was positive. Scheffé posthoc analyses revealed that the only significant difference between shows on any subscale was the difference between Show C (48%), for which 31% of classes had reported audio or video reception problems, and Show H (89%) on the Technical Reception Subscale.

5. SUMMARY AND CONCLUSIONS

On the following pages, the major findings of the evaluation of the eight-show "Infinity Factory" mini-series will be summarized.

Attention

Overall attention for eight shows was very high (mean = 91.3%). Attention rates were higher for Shows A and B, dropped to the lowest point for Show C, then stabilized near the 90% level for the remaining shows.

Despite significant differences found among the eight shows, posthoc analyses revealed that there were no significant differences between individual shows. In addition, no significant linear trend over the eight shows was found.

Student Appeal

Overall ratings of appeal on all subscales (show overall, "Brownstone" segments, "City Flats," "Scoops' Place," and animation segments) were very positive for all students, with all subscale ratings above 2.16 on a three-point rating scale.

A significant downward linear trend was reported for all subscales (show overall, "Brownstone" segments, "City Flats," "Scoops' Place," and animation segments). Although statistically significant, these declines were quite small (less than 5% over eight shows for "City Flats," "Scoops' Place," and animation segments, and less than 10% for the shows overall and "Brownstone" segments), and could be expected as the novelty of participating in a study and of watching a television show in school gradually wears off over a period of time.

Black and Latino students had higher appeal scores for the show overall. "Brownstone" segments. "City Flats," and "Scoops' Place" than did non-target students. Of particular interest is the finding that Black and Latino students' appeal ratings for "Scoops' Place" (featuring a Black cast) and "City Flats" (featuring a Latino cast) were comparable.

A significant two-way interaction of ethnic group and sex was reported for show overall, "Brownstone" segments, and "City Flats." This interaction indicated that Latino and non-target girls had greater appeal scores for these segments than Latino and non-target boys. In contrast, Black boys had greater appeal scores for these same segments than Black girls.

Appeal scores were comparable for younger and older students for the show overall, "Brownstone" segments, "Scoops' Place," and animation segments. However, appeal scores for "City Flats" were greater for younger students than for older students.

Appeal scores for the show overall, "Brownstone" segments, and "City Flats" were higher for girls than for boys while appeal ratings for "Scoops' Place" and animation segments were comparable for both sexes.

Comprehension of Dramatic Story Line

The overall mean percentage correct on the 24-item Comprehension Subscale (77.7%) was considered quite high. There were no differences among Comprehension Subscale scores for Black, Latino, and non-target students. Comprehension Subscale scores of older students were higher than those of younger students, as would be expected. Scores were comparable for boys and girls.

Knowledge of Math Content

Black, Latino, and non-target students performed

comparably on the Math Content Pretest. Scores were significantly higher for older than younger students, and significantly higher for boys than for girls.

All student groups showed significantly improved scores on the math content posttest after eight programs. The mean gain for all students on the Math Content Pretest/Posttest was 3.1; the mean score on the Math Content Subscale was 73.8%.

Both gain scores and Math Content Subscale scores were significantly greater for non-target than Black and Latino students. There was an interaction between ethnic group and sex gain scores: while gain scores were greater for Black and non-target girls than Black and non-target boys, Latino boys showed greater gains than Latino girls.

Both gain scores and subscale scores on knowledge of math content were comparable for boys and girls.

Of particular note are the findings that all student groups showed significant gains on both measures of knowledge of math content, and that non-target students benefitted substantially from watching these television programs which feature Black and Latino characters.

Student Attitudes

The overall mean gain on the 14-item pretest/posttest on social and math attitudes related to series goals was significant at the $p < .001$ level. Black students had significantly greater scores than non-target students, while there were no significant differences between Black and Latino students and between Latino and non-target students. There were no ethnic group differences found on the Attitude Subscale.

There were no significant age differences found between the gain scores on the 14-item social and math attitude measure. However, older students had significantly greater scores on the Attitude Subscale than younger students.

Girls had significantly greater scores than boys on the 14-item social and math attitudes measure. However, there were no significant sex differences on the Attitude Subscale.

There were no significant differences between overall mean pretest and posttest scores on measures of attitudes toward television and attitudes toward math. In addition, there were no significant differences among ethnic groups, age groups, and sexes on these measures.

The overall mean gain on the pretest/posttest measure of attitudes toward television programs about math was significant at the $p < .001$ level. Black students had significantly greater gain scores than non-target students, while there were no significant differences between Black and Latino and between Latino and non-target students. There were no significant age or sex differences found on this measure.

Of particular note are the findings that there were significant overall gains from pretest to posttest on measures of social and math attitudes related to series goals, and on attitudes toward television programs on math. On the whole, however, attitudinal gains from pretest to posttest were not as great as those found on knowledge of math content. In addition, findings were not as dramatically positive as those found on attention, appeal, comprehension, and teacher responses. This result may be due in part to the difficulty of changing attitudes and of measuring that change, after only eight programs.

Teacher Responses

There was a significant positive change in teacher attitudes toward educational television as measured by a semantic differential on that concept. This finding suggests that teachers found their experience with the "Infinity Factory" eight-show series to be a positive one.

Teachers' responses on the Teacher Weekly Questionnaire were generally positive; mean scores were above 70% for all subscales except the Class Preparation Sub-scale (mean = 43%). While there were significant differences among shows on six subscales (Presentation, Technical Reception, Student Attention, Program Appeal, Math Content, and Social Attitudes), when shows were compared pairwise on the difference between Show C (low) and Show H (high) on the Technical Reception Subscale was significant. In addition, the only significant linear trend over shows was a positive trend for the Technical Reception Subscale.

At the end of the eight-week evaluation period, 86% of the 36 teachers responding to the open-ended opinion form described their overall opinion of the series as positive. Specific aspects of the programs which received favorable comments were:

-- cultural and social aspects (42%);
-- effectiveness of the program in motivating students toward math and relating math to real life (44%);
-- cultural and social aspects of the program (42%);
-- student appeal and program presentation (33%); and
-- math content (28%).

While 28% commented favorably about the math content, 36% commented that there should be more math content or a fuller development of math content. In general, the level of math content was considered more appropriate by teachers of younger than older students.

Two factors should be considered in interpreting the teachers' responses to the programs, particularly the finding that in general teachers considered the programs

more effective in motivating students toward math and relating math to real life than they did in presenting math content. First, the programs are designed for home viewing as well as classroom use, and a lower density of math content may be necessary in order to compete for the home viewing audience. Second, the programs are designed principally to introduce math topics, show their relevance, and raise students' interest in math; they were not designed to be a complete instructional package. It is expected that teachers would be able to develop and carry out further instructional activities which would capitalize on the interest developed by the television programs and complete the learning process which the programs helped begin.

Conclusions

Results of the evaluation study support the conclusions that the "Infinity Factory" programs are able to capture and hold students' attention; that the programs have high appeal for Black, Latino, and non-target students; that the mathematical objectives of the series were generally met for all student groups; and that teachers consider the programs effective and useful.

Significant overall improvements were found in two of the four areas of students' attitudes measured (social and math attitudes related to series goals, and attitudes toward television programs about math). On the whole, however, results in the areas of student attitudes were not as dramatically positive as results in the areas of knowledge of math content, student attention, and student appeal.

6. RECOMMENDATIONS*

The "Infinity Factory" eight-show series was evaluated with over 1,000 students in Boston, Massachusetts; Lawrence, Massachusetts; New York, New York; and Los Angeles, California. The program series met with overall success as determined by measures of attention, appeal, comprehension, knowledge of math content, students' attitudes, and teacher responses. The following recommendations are made for distribution of the series, use in schools, ongoing production, and ongoing evaluation.

1. Dissemination efforts should stress the advantages of the series for a non-target audience as well as the target audience since evaluation findings indicate that non-target students liked the program and benefitted from math content presented in a multicultural context.
2. Dissemination efforts aimed toward schools should include more extensive orientation for teachers, with particular emphasis on the series' objectives (both math and cultural/social), the rationale for the series, and how the television programs and program guides can be used more effectively in schools.
3. The program guides should include more ways to develop positive social, cultural, and math attitudes in a classroom setting since many teachers rated the program very high in these areas but focussed mainly on math content in related classroom activities they conducted.

---

*These recommendations are based on the findings of both Part I and Part II of the evaluation study, and on the experiences of the evaluators in conducting the study. The recommendations are also included as Section 13 of Part I.

4. In ongoing production, special attention should be given to defining both math and cultural/social program objectives more clearly, and to carefully translating these objectives into program content.
5. Since the historical and "Math in the Street" segments did not hold attention over the eight-show series, alternative presentation of the content in these sequences should be considered. Also, certain "Math Fact" segments were found to have low appeal for students, and alternatives for these should be considered.
6. Evaluators found that show segments which feature Black or Latino families are well received by both Black and Latino students. Therefore, this type of approach to multicultural education should be continued in future productions.
7. Since measures of appeal and comprehension of story line were quite high for the target audience and math content measures somewhat lower, it is suggested that one possible route to improving students' math learning might be a better integration of dramatic story line and math content.
8. Special attention should be paid to the diction of cast members. It is recommended that the series use only actors who can be clearly understood by all segments of the target audience.
9. The evaluation findings indicate that younger and older students respond differently to the programs, especially in the areas of math content, attitudes, and comprehension of story line. Therefore, special attention should be given to the development of material appropriate for specific age groups, and each program should contain

material appropriate for different age levels of the target audience.

10. The purpose of some shorter segments of the shows was not clear to the evaluators. It is recommended that the educational objectives, both math and cultural/social, of these shorter segments be planned and integrated into the context of the overall program as carefully as the major segments.
11. Ongoing content analysis of programs should be implemented during production in order to monitor the degree to which each program segment meets its objectives. This information would also be useful to evaluators in developing items for criterion-referenced measures of program impact.
12. More careful attention should be given to relating programs and program segments to specific cultural and social objectives of the series.
13. Although the eight-show series was evaluated in the schools, the programs were designed primarily for home-viewing. Ongoing formative evaluation efforts should attempt to examine program impact in a non-school setting.
14. A further examination of the effects of attention and appeal on comprehension, math attitudes, and knowledge of math content is recommended.
15. Due to the cumbersome and time-consuming nature of a large-scale evaluation effort, more informal evaluation efforts are recommended with smaller groups of children. It is further recommended that this process be built into the program development process from the planning stages on, in order to provide ongoing and more immediate feedback throughout all phases of program development.
16. Greater time should be allowed for the development of evaluation measures so that more exten-

sive piloting procedures could be initiated. The smaller-scale, ongoing evaluation recommended above would facilitate this process. Special attention should be given to the development of measures of appeal and attitudes, which are especially difficult to measure.